Kybernetika

Guang-Da Hu
A modified version of explicit Runge-Kutta methods for energy-preserving

Kybernetika, Vol. 50 (2014), No. 5, 838–847

Persistent URL: http://dml.cz/dmlcz/144110

Terms of use:

© Institute of Information Theory and Automation AS CR, 2014

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

KYBERNETIKA — VOLUME 50 (2014), NUMBER 5, PAGES 838–847

A MODIFIED VERSION OF EXPLICIT RUNGE–KUTTA METHODS FOR ENERGY-PRESERVING

Guang-Da Hu

In this paper, Runge–Kutta methods are discussed for numerical solutions of conservative systems. For the energy of conservative systems being as close to the initial energy as possible, a modified version of explicit Runge–Kutta methods is presented. The order of the modified Runge–Kutta method is the same as the standard Runge–Kutta method, but it is superior in energy-preserving to the standard one. Comparing the modified Runge–Kutta method with the standard Runge–Kutta method, numerical experiments are provided to illustrate the effectiveness of the modified Runge–Kutta method.

Keywords: energy-preserving, explicit Runge–Kutta methods, gradient

Classification: 65L05, 65L07

1. INTRODUCTION

In this paper, we consider the numerical solutions of the system of nonlinear differential equations

$$\dot{x}(t) = f(x(t)), \tag{1}$$

where $f : R^d \to R^d$ is continuously differentiable, $x(t) \in R^d$ is state variable vector and the initial vector $x(0) = x_0$ is known. We always assume that system (1) has a unique solution. Furthermore, assume that an energy function of system (1), $E(x(t))$ is a known constant

$$E(x(t)) = E(x(0)), \tag{2}$$

i. e., the energy $E(x(t))$ is conservative for any $t \geq 0$.

System (1) has the first integral $E(x(t))$, and numerical methods that preserve this integral are usually called energy-preserving methods. The energy-preserving method is one of the geometric numerical integration methods [7, 8, 10].

All Runge–Kutta (RK) methods preserve arbitrary linear invariants [7, 12], and some (the symplectic) RK methods preserve arbitrary quadratic invariants [5]. However, no RK methods can preserve arbitrary polynomial invariants of degree 3 or higher [4].

In this paper, we are only concerned with explicit RK methods. We emphasize that we do not require $E(x_n) = E(x(0))$ exactly holds, but approximately holds when explicit RK methods are applied to general conservative system (1) with (2). We study numerical

DOI: 10.14736/kyb-2014-5-0838

methods that provide an approximation to the solution of (1) with (2) preserving the energy in the sense that $E(x_n)$ approximates to $E(x(0))$.

The conservative systems appear in many problems in science and engineering, for instance, mechanical vibration, celestial mechanics and Hamiltonian systems [11].

In order to numerically preserve the energy of conservative systems, many numerical methods have been developed [1, 2, 3, 8]. Because the numerical methods require $E(x_n) = E(x(0))$ holds exactly for numerical solutions, only implicit numerical methods are possible for general conservative system (1). Projection methods are developed for the numerical solutions of differential equations on manifolds [3, 8]. The methods are also applied to conservative system (1) with (2). The basic principle of the projection RK methods for preserving-energy is as follows. An arbitrary RK method is applied to system (1) to obtain a numerical solution, then project the numerical solution onto the manifold which is determined by the energy function. The methods require to solve a system of nonlinear equations at each step. The projection method is implicit even if an explicit RK scheme is applied [3, 8]. Hence the computational effort is large to implement projection RK methods. Recently the line integral methods are introduced in [2] and developed in [1] which have been applied to conservative systems. The key idea of these methods are imposing energy conservation through a line integral. These methods are also implicit. An explicit RK scheme is presented in [6] which preserves quadratic conservation laws. We have to point out that the explicit RK scheme in [6] can give the exact values of quadratic functions. The idea of the explicit RK scheme in [6] is interesting.

With respect to the computational cost, explicit RK methods are superior to implicit ones. However, the standard explicit RK methods do not preserve the energy function well. For general conservative system (1) with (2), it is impossible for any explicit RK methods to exactly preserve the energy [4]. When explicit RK methods are applied to general conservative system (1) with (2), the requirement $E(x_n) = E(x(0))$ must be relaxed to develop explicit numerical methods. We only require that $E(x_n)$ is as close to $E(x(0))$ as possible. The aim of numerical computation is to seek an approximated (numerical) solution which is as close to the exact solution as possible. Because a numerical solution is approximated, it is not necessary for the energy function of the numerical solution must be preserved exactly. We require that the energy function is also approximated. Hence it is reasonable that the energy function of the numerical solution is as close to the exact energy as possible. In this paper, based on the idea, with the gradient of the energy function a modified version of explicit RK methods is presented.

This paper is organized as follows. A modified revision of explicit RK methods is presented in section 2. In section 3, comparing the modified revision of the explicit RK methods with standard RK method, numerical experiments are provided to illustrate the effectiveness of the modified RK method. Conclusion is given in section 4.

2. A MODIFIED VERSION OF EXPLICIT RUNGE–KUTTA METHODS

In this section, considering system (1) with (2), we present a modified version of explicit RK methods with the gradient of energy function such that the energy of the numerical solution is as close to the initial energy as possible.

The gradient of $E(x)$ is defined by

$$\nabla E(x) = \left[\frac{\partial E(x)}{\partial x_1}, \frac{\partial E(x)}{\partial x_2}, \ldots, \frac{\partial E(x)}{\partial x_d}\right]^T,$$

where $x = [x_1, x_2, \ldots, x_d]^T$.

Now we consider the numerical solution of system (1) by RK methods. Let x_n denote an approximation (numerical solution) to the solution $x(t_n)$ of system (1) at t_n, i. e., $x_n \approx x(t_n)$. The standard $s-$stage RK method for system (1) is defined by the difference system [7]

$$\begin{cases} x_{n+1} = x_n + h\sum_{i=1}^{s} b_i g_i, \\ g_i = f(x_n + h\sum_{j=1}^{s} a_{ij} g_j), \qquad i = 1, 2, \ldots, s. \end{cases} \tag{3}$$

Here h stands for a step-size and $t_n = t_{n-1} + h$. When $a_{ij} = 0$ for $i \le j$, the RK method is explicit. Hereafter we will only consider explicit RK methods.

Assume that the continuous-time system (1) has a known energy function $E(x(t))$ with $E(x(t)) = E(x(0))$. We want to seek a modified version of explicit RK methods which produces a numerical solution x_n such that $E(x_n)$ is as close to $E(x(0))$ as possible. In addition, the order of the modified version is the same as the standard RK methods. Since we are only concerned with explicit RK methods, it is emphasized that we only require that $E(x_n) = E(x(0))$ approximately holds.

Now we present a modified version of the pth order explicit RK method (3) with the modified formula

$$\bar{x}_{n+1} = \begin{cases} x_{n+1} - h^{p+1}\nabla E(x_{n+1}), & \text{if} \quad \triangle E(x_{n+1}) > 0, \\ x_{n+1}, & \text{if} \quad \triangle E(x_{n+1}) = 0, \\ x_{n+1} + h^{p+1}\nabla E(x_{n+1}), & \text{if} \quad \triangle E(x_{n+1}) < 0, \end{cases} \tag{4}$$

where $\triangle E(x_{n+1}) = E(x_{n+1}) - E(x(0))$.

Now we present the main result of this paper.

Theorem 2.1. Assume that the localizing assumption $x_n = x(t_n)$ holds, the energy function $E(x(t))$ of nonlinear system (1) is conserved and continuously differentiable. Let the modified RK method (3) – (4) be applied to system (1) with (2). If the step-size h is sufficiently small and the gradient $\nabla E(x_{n+1}) \neq 0$, then

$$E(\bar{x}_{n+1}) < E(x_{n+1}), \quad \text{for} \quad \triangle E(x_{n+1}) > 0 \tag{5}$$

and

$$E(\bar{x}_{n+1}) > E(x_{n+1}), \quad \text{for} \quad \triangle E(x_{n+1}) < 0 \tag{6}$$

hold. Furthermore, the order of the modified version (3) – (4) is the same as the standard one (3).

P r o o f. For $\triangle E(x_{n+1}) > 0$, by formula (4),

$$\bar{x}_{n+1} = x_{n+1} - h^{p+1}\nabla E(x_{n+1}). \tag{7}$$

According to Taylor expansion of $E(\bar{x}_{n+1})$ at x_{n+1} and (7), we have

$$E(\bar{x}_{n+1}) = E(x_{n+1}) - [\nabla E(x_{n+1})]^T [\nabla E(x_{n+1})] h^{p+1} + O(h^{2(p+1)}). \tag{8}$$

Since $\nabla E(x_{n+1}) \neq 0$ and h is sufficiently small, we obtain

$$E(\bar{x}_{n+1}) < E(x_{n+1}).$$

Similarly, for $\triangle E(x_{n+1}) < 0$, by formula (4),

$$\bar{x}_{n+1} = x_{n+1} + h^{p+1} \nabla E(x_{n+1}), \tag{9}$$

we obtain

$$E(\bar{x}_{n+1}) > E(x_{n+1}).$$

Since the energy function $E(x(t))$ is continuously differentiable, $||\nabla E(x_{n+1})||$ is bounded. By the modified formula (4), $\bar{x}_{n+1}$ has the same order p as x_{n+1}. The proof is completed. □

Remark 2.1. Formula (4) may be extended to the following form

$$\bar{x}_{n+1} = x_{n+1} + \lambda h^{p+1} \nabla E(x_{n+1}), \tag{10}$$

where the parameter λ is a real number. An optimal choice of the parameter λ can be done by

$$\min_{\lambda} [E(\bar{x}_{n+1}) - E(x(0))]^2.$$

It involves solving a line (one-dimensional) search.

Based on the modified RK method (3) with (10), we propose the following computational procedure to numerically solve system (1) with (2).

Algorithm

Step 1. Take the initial value x_0, the number of steps N, the step-size h and $n = 0$.

Step 2. Check whether $n = N$. If $n = N$, then end the procedure; otherwise, compute x_{n+1} by (3).

Step 3. Seek the parameter λ such that

$$\min_{\lambda} [E(\bar{x}_{n+1}) - E(x(0))]^2,$$

where

$$\bar{x}_{n+1} = x_{n+1} + \lambda h^{p+1} \nabla E(x_{n+1}),$$

which is from (10). Then, let $x_{n+1} = \bar{x}_{n+1}, n = n + 1$, and go to Step 2.

By Step 1 to Step 3, one can obtain the numerical solution $\bar{x}_n, n = 1, 2, \ldots, N$.

Remark 2.2. We have to point out that Theorem 2.1 is only a local property of the modified formula. It is a further topic to study the efficiency of the above algorithm in general.

It is obvious to obtain the following result.

Theorem 2.2. Assume that the conditions of Theorem 2.1 are satisfied. For the numerical solution $\bar{x}_{n+1}$ produced by the above procedure, the inequality

$$|\triangle E(\bar{x}_{n+1})| \leq |\triangle E(x_{n+1})| \tag{11}$$

holds, where x_{n+1} is obtained by the standard RK method (3) and $\triangle E(x_{n+1}) = E(x_{n+1}) - E(x(0))$. Furthermore, the order of the above procedure is the same as the standard RK method (3).

Remark 2.3. Since the above procedure is explicit, it is superior in computational cost to the project methods in [8] and the line integral methods in [1, 2] which are implicit. Inequality (11) in Theorem 2.2 shows that the modified RK method (3) with (10) for system (1) is superior in energy-preserving to the standard one (3).

Remark 2.4. Assume that system (1) has m first integrals, for $i = 1, 2, \ldots, m$, $E_i(x(t)) = E_i(x(0)) = c_i$. Let

$$E(x(t)) = \sum_{i=1}^{m} [E_i(x(t)) - c_i]^2.$$

Then the above algorithm can be applied to system (1) with m first integrals. It is a further research topic.

3. NUMERICAL EXPERIMENTS

In this section, two typical conservative systems are given to demonstrate the effectiveness of the procedure derived based on the modified RK method (3) with (10). The modified version of the classical fourth-order explicit RK method [7] will be used, i. e.,

$$\begin{cases} x_{n+1} = & x_n + \frac{h}{6}(g_1 + 2g_2 + 2g_3 + g_4) \\ g_1 = & f(x_n) \\ g_2 = & f(x_n + \frac{h}{2}g_1) \\ g_3 = & f(x_n + \frac{h}{2}g_2) \\ g_4 = & f(x_n + hg_3) \end{cases} \tag{12}$$

with the modified formula

$$\bar{x}_{n+1} = x_{n+1} + \lambda h^{p+1} \nabla E(x_{n+1}). \tag{13}$$

We will compare the modified version (12) – (13) with the fourth-order classical RK method (12). The phase diagrams of numerical solutions and the corresponding energy functions produced by the two methods will be shown, respectively. All the computations are carried out by means of Matlab 7.0.

Example 3.1. Consider the simple pendulum without friction described by the nonlinear differential equation [9]

$$\ddot{x} + a \sin x = 0 \tag{14}$$

where $a > 0$ is constant, and $x(t)$ measures the angle formed by the pendulum and the vertical downward direction.

To change the second-order equation (14) into a first-order system, we let $y = \dot{x}$. Then obtain the system

$$\begin{cases} \dot{x} = y \\ \dot{y} = -a \sin x. \end{cases} \tag{15}$$

We take $a = 1$, the step-size $h = 0.1$ and the initial position $[x(0), y(0)]^T = [\frac{2\pi}{3}, 0]^T$.

For system (15), the energy function [9]

$$E(x, y) = \frac{1}{2} y^2 + 1 - \cos x.$$

The derivative of $E(x, y)$ is

$$\dot{E}(x, y) = 0.$$

It means that the trajectories $x(t), y(t)$ starting on the energy surface $E(x(0), y(0))$ remain on the same surface for all future time.

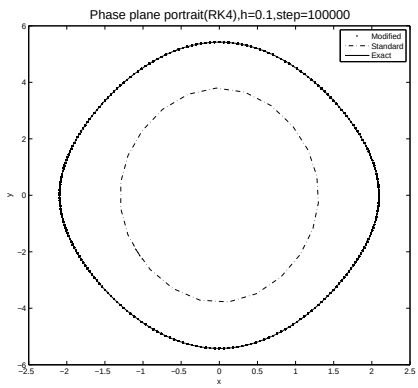

Fig. 1. The phase portraits of $y(t)$ and $\dot{y}(t)$.

Figure 1 shows that the phase portraits of $x - y$ by the modified RK method and standard one, respectively when the time step is near to 100000. They are denoted by $\cdot$ line and dotted line, respectively. The exact phase portrait is denoted by solid line.

Figure 2 shows that the energy functions of numerical solutions by the modified RK method and standard one, and are also denoted by $\cdot$ line and dotted line, respectively. The exact energy is denoted by solid line.

For comparison, the values obtained using the modified RK method and the standard RK are shown. Both the methods produce an accurate solution for small values of t, but

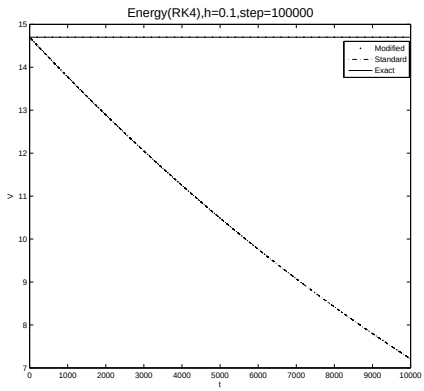

Fig. 2. Energy function $E(t)$.

for large values the standard RK method differs significantly from the modified version. For this example, the energy decay using the standard RK method is substantial, whereas the modified version produces a value of the energy function that oscillates but remains very near the exact value over the entire time period, see Figure 2. It is also apparent in Figure 1 that the position and the velocity obtained from the modified version, even for large values of t, are very nearly on the curve followed by the exact solution, whereas the standard RK method provides very poor approximations.

For 100000 time steps, the number of updates of the modified RK method for this example is 126643. The average number of updates in each time step is 1.26643.

Remark 3.1. In example 1, the energy function is not quadratic form, the scheme in [6] can not work.

Example 3.2. Consider a Hamiltonian system described by the nonlinear differential equations [11]

$$\begin{cases} \dot{q} = \frac{\partial H(q,p)}{\partial p} \\ \dot{p} = -\frac{\partial H(q,p)}{\partial q} \end{cases} \tag{16}$$

where $q = [q_1, q_2]^T$ and $p = [p_1, p_2]^T$ are the generalized position and generalized momentum coordinates, respectively. The Hamiltonian function is

$$H(q,p) = \frac{1}{2m}[(p_1)^2 + (p_2 - kq_1)^2]. \tag{17}$$

The energy function $E(q,p) = H(q,p)$ for the Hamiltonian system.

The parameters of the Hamiltonian system are $m = 1$ and $k = 1$. Assuming that the initial conditions are $q_1(0) = -1$ $q_2(0) = 2$, $p_1(0) = 0$ and $p_2(0) = D = 4$, then the exact

solution $q = [q_1, q_2]^T$ satisfies the circle equation

$$\left(q_1 - \frac{D}{k}\right)^2 + (q_2 - C)^2 = \frac{2mE_0}{k^2},$$

where

$$C = q_2 - \frac{1}{k}\sqrt{2mE_0 - (D - kq_1)^2},$$

and

$$E_0 = E(q(0), p(0)).$$

Taking $h = 0.2$, the results of the calculation are shown in Figures 3 and 4.

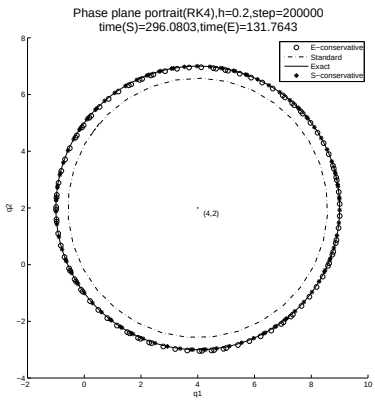

Fig. 3. The phase portraits of $q_1(t)$ and $q_2(t)$.

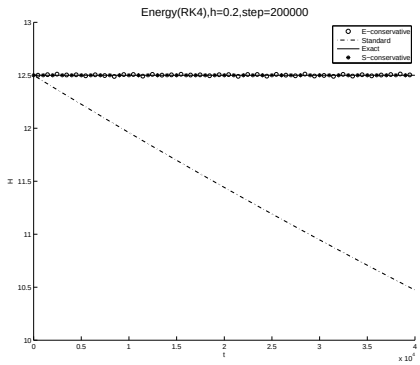

Fig. 4. Energy function $E(t)$.

Figure 3 shows that the phase portraits of $q_1 - q_2$ by the modified RK method, the standard one and the scheme in [6], respectively when the time step is near to 200000. They are denoted by $\circ$ line, dotted line and $*$ line, respectively. The exact phase portrait is denoted by solid line.

Figure 4 shows that the energy functions of numerical solutions by the modified RK method, the standard one, and the scheme in [6], respectively. They are also denoted by $\circ$ line, dotted line and $*$ line, respectively. The exact energy is denoted by solid line.

For comparison, the values obtained using the modified RK method and the standard RK are shown. Both the methods produce an accurate solution for small values of t, but for large values the standard method differs significantly from the modified version. For this example, the energy decay using the standard RK is substantial, whereas the modified version produces a value of E that oscillates but remains very near the exact value over the entire time period, see Figure 4. It is also apparent in Figure 3 that the generalized positions obtained from the modified version, even for large values of t, are very nearly on the circle followed by the exact solution, whereas the standard RK method provides very poor approximations.

For 200000 time steps, the number of updates of the modified RK method for this example is 200727. The average number of updates in each time step is 1.0036.

Since the energy function in this example is quadratic form, we can compare the modified RK method with the scheme in [6]. For this example, we have the following conclusion.

Remark 3.2. Both the modified RK method and the explicit scheme in [6] can work well, see Figures 3 and 4. The computation time of the modified RK and the explicit scheme in [6] are 131.7643 s and 296.0803s, respectively. It shows that the modified RK method is faster than the scheme in [6]. We have to point out that the explicit scheme in [6] can give the exact values of quadratic functions and the modified RK method (3) with (10) can not give the exact values.

Remark 3.3. The above two examples show that the numerical solutions produced by the procedure based on the modified RK method (3) with (10) is superior to the standard RK method (3).

4. CONCLUSION

In this paper, a modified version of explicit RK methods is presented for numerical solutions of conservative systems. A procedure based on the modified RK method is superior in energy-preserving to the standard RK method. Since the modified version of RK methods is explicit, it is easier to implement than the implicit methods in literature for energy-preserving. Numerical experiments show the effectiveness of the procedure derived based on the modified RK method.

ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (11371053 and 61372090).

The author wishes to thank the two anonymous reviewers for their valuable comments and suggestions which are helpful to improve this paper. One of the reviewers let the author know the reference [6] and revise the algorithm.

(Received June 26, 2013)

REFERENCES

[1] L. Brugnano, M. Calvo, J. I. Montijano, and L. Rândez: Energy-preserving methods for Poisson systems. J. Comput. Appl. Math. *236* (2012), 3890–3904.

[2] L. Brugnano, F. Iavernaro, and D. Trigiante: Hamiltonian boundary value methods (energy-preserving discrete line integral methods). J. Numer. Anal. Ind. Appl. Math. *5* (2010), 1–2, 17–37.

[3] M. Calvo, M. P. Laburta, J. I. Montijano, and L. Rândez: Error growth in numerical integration of periodic orbits. Math. Comput. Simul. *81* (2011), 2646–2661.

[4] M. Calvo, A. Iserles, and A. Zanna: Numerical solution of isospectral flows. Math. Comput. *66* (1997), 1461–1486.

[5] G. J. Cooper: Stability of Runge–Kutta methods for trajectory problems. IMA J. Numer. Anal. *7* (1987), 1–13.

[6] N. Del Buono and C. Mastroserio: Explicit methods based on a class of four stage Runge–Kutta methods for preserving quadratic laws. J. Comput. Appl. Math. *140* (2002), 231–243.

[7] D. F. Griffiths and D. J. Higham: Numerical Methods for Ordinary Differential Equations. Springer-Verlag, London 2010.

[8] E. Hairer, C. Lubich, and G. Wanner: Geometric Numerical Integration. Springer-Verlag, Berlin 2002.

[9] H. K. Khalil: Nonlinear Systems. Third Edition. Prentice Hall, Upper Saddle River, NJ 2002.

[10] T. Lee, M. Leok, and N. H. McClamroch: Lie variational integrators for the full body problem in orbital methanics. Celest. Meth. Dyn. Astr. *98* (2007), 121–144.

[11] S. Li: Introduction to Classical Mechanics. (In Chinese.) University of Science and Technology of China, Hefei 2007.

[12] L. F. Shampine: Conservation laws and numerical solution of ODEs. Comput. Math. Appl. *12B* (1986), 1287–1296.

Guang-Da Hu, School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, 100083. P. R. China.

e-mail: ghu@hit.edu.cn